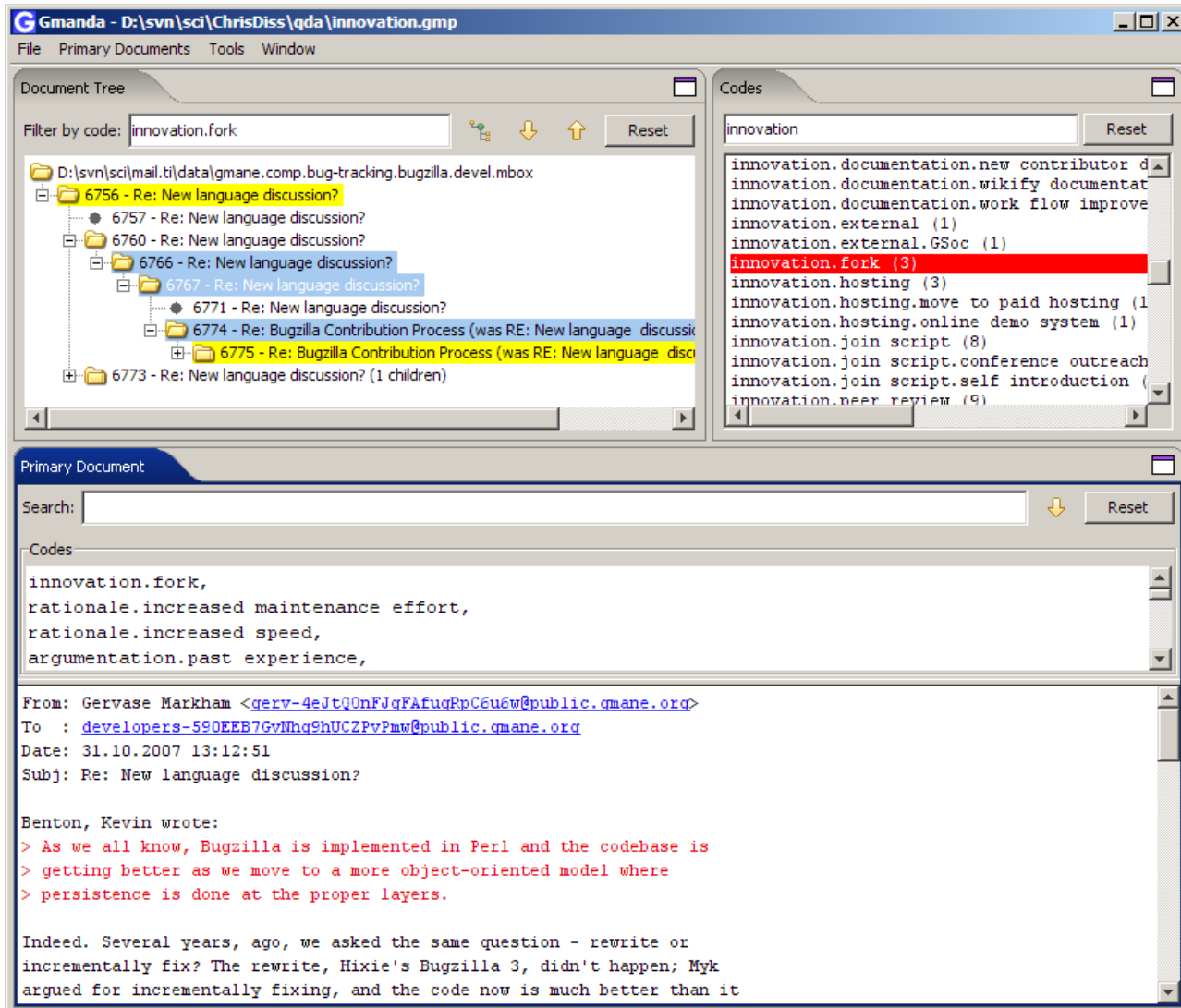




# **Research Ethics for Studying Open Source Projects**

Christopher Oezbek  
Freie Universität Berlin  
23.02.2008



The screenshot displays the Gmanda application window. The title bar reads "Gmanda - D:\svn\sci\ChrisDiss\qda\innovation.gmp". The interface is divided into several panes:

- Document Tree:** Shows a hierarchical view of a document. The root is "D:\svn\sci\mail.ti\data\gmane.comp.bug-tracking.bugzilla.devel.mbox". Underneath, there are several folders representing email threads, such as "6756 - Re: New language discussion?". The folder "6775 - Re: Bugzilla Contribution Process (was RE: New language discussion)" is highlighted in yellow.
- Codes:** A list of codes used for coding. The code "innovation.fork (3)" is highlighted in red. Other codes include "innovation.documentation.new contributor d", "innovation.documentation.wikify documentat", "innovation.documentation.work flow improve", "innovation.external (1)", "innovation.external.GSoc (1)", "innovation.hosting (3)", "innovation.hosting.move to paid hosting (1)", "innovation.hosting.online demo system (1)", "innovation.join script (8)", "innovation.join script.conference outreach", "innovation.join script.self introduction (", and "innovation.near\_review (9)".
- Primary Document:** Shows the content of the selected document. It includes a search bar, a "Codes" section with the following text:

```
innovation.fork,  
rationale.increased maintenance effort,  
rationale.increased speed,  
argumentation.past experience,
```

and an email header:

```
From: Gervase Markham <gerv-4eJtQQnFJqFAfugPpCGu6w@public.gmane.org>  
To : developers-590EEB7GvNhg9hUCZPvPmw@public.gmane.org  
Date: 31.10.2007 13:12:51  
Subj: Re: New language discussion?
```

Below the header, it says "Benton, Kevin wrote:" followed by a quoted text block:

```
> As we all know, Bugzilla is implemented in Perl and the codebase is  
> getting better as we move to a more object-oriented model where  
> persistence is done at the proper layers.
```

At the bottom, there is a paragraph of text:

```
Indeed. Several years, ago, we asked the same question - rewrite or  
incrementally fix? The rewrite, Hixie's Bugzilla 3, didn't happen; Myk  
argued for incrementally fixing, and the code now is much better than it
```

Maintainer to committer:

*„...next time I see crap like this I'm going to...  
[...]  
I'm going to go postal on the next  
maintainer who doesn't understand what "merge  
window" and "fixes only" means.“*



Can I publish this?



"next time I see crap like this I'm going to"

Search

[Advanced Search](#)  
[Preferences](#)

Web

Results 1 - 1 of 1 for "[next time I see crap like this I'm going to](#)". (0.05 seconds)

[Re: \[GIT PULL\]\[RETRY\] KVM updates for Linux 2.6.21-rc2 \[LWN.net\]](#)

6 Mar 2007 ... I want \*FIXES\*ONLY\* by now, and **next time I see crap like this I'm going to ignore it until you fix me a pull that contains fixes only, ...**

[lwn.net/Articles/224886/](#) - 13k - [Cached](#) - [Similar pages](#)

From a paper about introducing an information manager role at GNU Classpath:

*„I do not like wikis.“, which is not further elaborated upon by the author.“*

Google™

classpath "I do not like wikis."

Search

[Advanced Search](#)  
[Preferences](#)

Web

Your search - **classpath "I do not like wikis."** - did not match any documents.



classpath "I hate wikis."

Search

[Advanced Search](#)  
[Preferences](#)

Web

Results 1 - 5 of 5 for classpath "[I hate wikis.](#)". (0.37 seconds)

[Re: Get ready for FOSDEM 2006](#)

**classpath.** <-- Thread -->; <-- Date --> ... **I hate Wikis.** OK, I'll do it. :-( Andrew. Get ready for FOSDEM 2006 Mark Wielaard ...

[www.mail-archive.com/classpath@gnu.org/msg11949.html](http://www.mail-archive.com/classpath@gnu.org/msg11949.html) - 6k - [Cached](#) - [Similar pages](#)

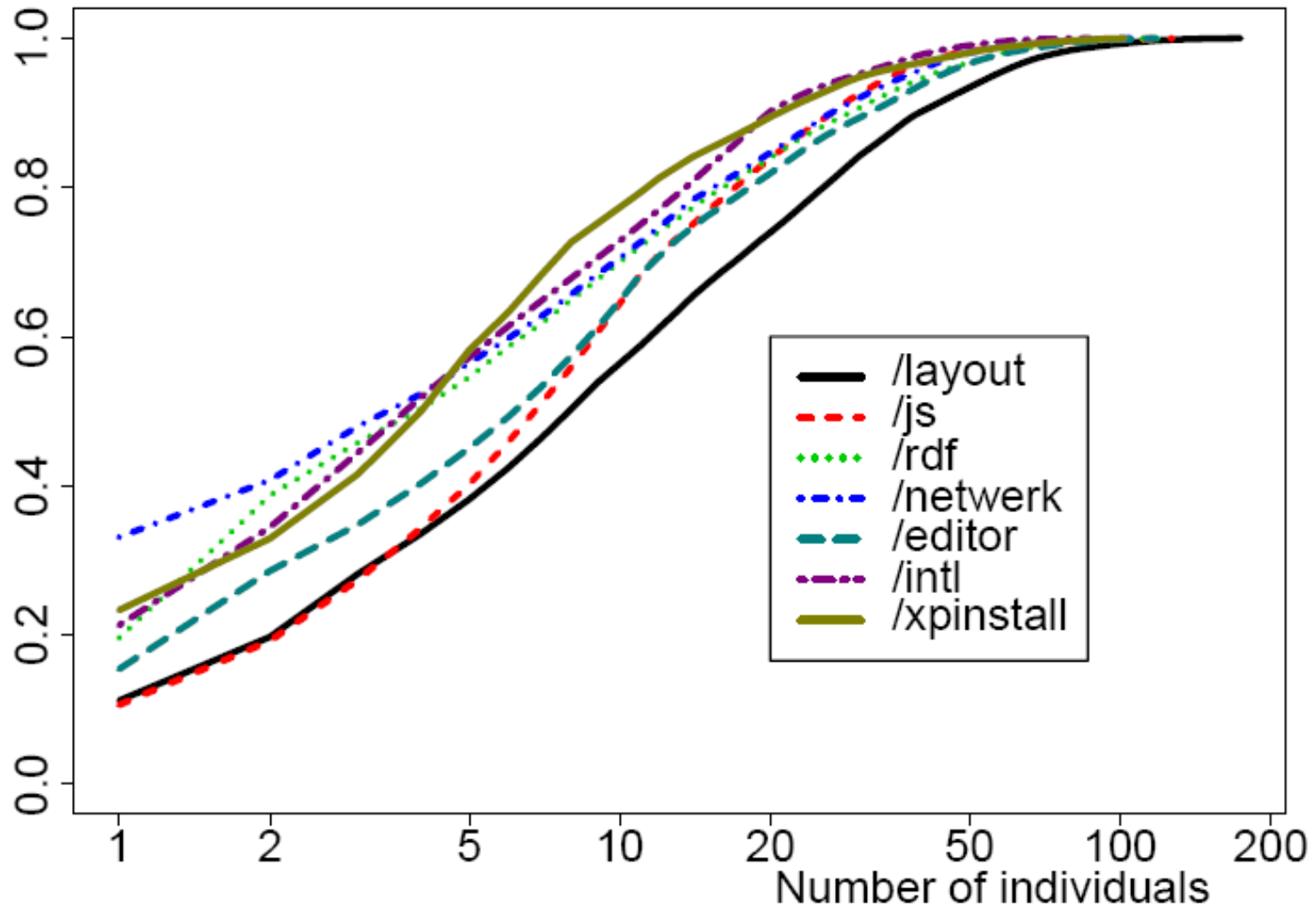
If individuals can be identified easily from research results...



Is there an ethical problem?



How about the other research methods in Open Source?



Mockus et al. 2002: Two Case Studies of Open Source Software Development: Apache and Mozilla

**Table 2: Respondent's primary open source projects**

<b>Multiple occur.</b>	<b>Single occurrences</b>			
Linux (21)	AbiWord	Genes	LTSP	Tabindex
Midgard (4)	Analog	Gtk Explorer	MPLS for Linux	The COG Engine
Perl (3)	Cons	Hover Carnage	NetBSD	Vaxbb

[Hars and Ou, 2001: Working for Free? – Motivations of Participating in Open Source Projects]

# Respect for Persons

Justice

Benefice



Discussion Point:  
What kind of data is public?



Discussion Point:  
What might be damaging to  
participants?



Discussion Point:  
Ethics review at your university?



Discussion Point:  
What can we / must we do?

# Thank you!

- What kind of data is public?
- What might be damaging to participants?
- Ethics review at your university?
- What can we / must we do?